JB
Journals.ASM.org

In Vitro Recombinants of Antibiotic-Resistant *Chlamydia trachomatis* Strains Have Statistically More Breakpoints than Clinical Recombinants for the Same Sequenced Loci and Exhibit Selection at Unexpected Loci

Tara Srinivasan,[a] William J. Bruno,[b] Raymond Wan,[a] Albert Yen,[a] Jennifer Duong,[a] and Deborah Dean[a,c,d]

Center for Immunobiology and Vaccine Development, Children's Hospital Oakland Research Institute, Oakland, California, USA[a]; T6 Division, Los Alamos National Lab, Los Alamos, New Mexico, USA[b]; University of California at Berkeley and University of California at San Francisco Graduate Program in Bioengineering, Berkeley, California, USA[c]; and Department of Medicine, University of California, San Francisco, San Francisco, California, USA[d]

Lateral gene transfer (LGT) is essential for generating between-strain genomic recombinants of *Chlamydia trachomatis* to facilitate the organism's evolution. Because there is no reliable laboratory-based gene transfer system for *C. trachomatis*, *in vitro* generation of recombinants from antibiotic-resistant strains is being used to study LGT. However, selection pressures imposed on *in vitro* recombinants likely affect statistical properties of recombination relative to naturally occurring clinical recombinants, including prevalence at particular loci. We examined multiple loci for 16 *in vitro*-derived recombinants of ofloxacin- and rifampin-resistant L_1 and D strains, respectively, grown with both antibiotics, and compared these with the same sequenced loci among 11 clinical recombinants. Breakpoints and recombination frequency were examined using phylogenetics, bioinformatics, and statistics. *In vitro* and clinical isolates clustered perfectly into two groups, without misclassification, using Ward's minimum variance based on breakpoint data. As expected, *gyrA* (confers ofloxacin resistance) and *rpoB* (confers rifampin resistance) had significantly more breakpoints among *in vitro* recombinants than among clinical recombinants ($P < 0.0001$ and $P = 0.02$, respectively, using the Wilcoxon rank sum test). Unexpectedly, *trpA* also had significantly more breakpoints for *in vitro* recombinants ($P < 0.0001$). There was also significant selection at other loci. The strongest bias was for *ompA* in strain D ($P = 3.3 \times 10^{-8}$). Our results indicate that the *in vitro* model differs statistically from natural recombination events. Additional genomic studies are needed to determine the factors responsible for the observed selection biases at unexpected loci and whether these are important for LGT to inform approaches for genetically manipulating *C. trachomatis*.

Chlamydia trachomatis is the most prevalent sexually transmitted bacterium and the leading cause of preventable blindness in the world. More than 92 million cases are reported annually (42). Infection leads to diseases such as urethritis, cervicitis, pelvic inflammatory disease, infertility, proctitis, reactive arthritis, lymphogranuloma venereum, and trachoma. Over 70% of women and 50% of men who are infected with *C. trachomatis* are asymptomatic, providing a huge opportunity for unchecked transmission (23, 42). For symptomatic individuals who do seek medical care, diagnostic tests for *C. trachomatis*, either for initial diagnosis or test of cure, are not always performed, largely due to cost (8). Thus, empirical treatment with antibiotics is common, which may hasten the development of drug resistance. This has already occurred for *Chlamydia suis*, a species that is closely related to *C. trachomatis* in which a tetracycline transposon was acquired from other pathogens residing in the guts of pigs (13). *C. trachomatis* may not be far behind; a number of studies have shown increased MICs of various antibiotics for clinical *C. trachomatis* isolates or reference strains propagated in tissue culture under selective antimicrobial pressure (2, 3, 26, 33–35).

In addition to the example of lateral gene transfer (LGT) in *C. suis*, a growing body of literature points toward recombination as a major mechanism by which *C. trachomatis* has advanced its evolutionary reproductive fitness and ability to evade the host immune response (14, 16–18, 20, 21, 25, 37). It is well accepted that, among asexual populations evolving in a changing environment, deleterious mutations increase without bound until they reach an intolerable level that would presumably cause population extinction, referred to as Muller's ratchet, unless rescued by recombination (27, 28). Agent-based simulations of finite populations have shown good agreement with analytical theory, where an increased mutation rate is strongly correlated with decreased fitness to the point of population collapse in the absence of recombination (15). These considerations explain why chlamydiae must include recombination, despite their intracellular lifestyle, in order to survive.

The study of the biology of chlamydiae is complicated by their obligate intracellular development and the lack of a routine laboratory system for directed mutagenesis or reliable gene transfer. However, knowledge of the mechanism(s) that chlamydiae utilize to undergo recombination would provide the necessary insight to presumably effectively genetically manipulate the organism. Toward this goal, investigators have developed an *in vitro* system to model *in vivo* LGT in chlamydiae. DeMars et al. were the first to culture different antibiotic-resistant D (rifampin) and L_1 (ofloxa-

Received 11 October 2011 **Accepted** 13 November 2011
Published ahead of print 28 November 2011
Address correspondence to Deborah Dean, ddean@chori.org.
Supplemental material for this article may be found at http://jb.asm.org/.

doi:10.1128/JB.06268-11

cin) strains of *C. trachomatis* to produce recombinant progeny that were doubly resistant when grown in media containing both antibiotics (11). Suchland et al. subsequently produced recombinants of various antibiotic-resistant strains of *C. trachomatis*, *C. suis*, and *Chlamydia muridarum*, including cross-species recombinant events (38). It is not clear whether recombinants created by antimicrobial pressure have the same breakpoint location and frequency of LGT events, or even necessarily the same mechanisms of recombination, compared with those circulating among human populations. Therefore, we examined multiple genomic loci for 16 *in vitro*-derived recombinants of ofloxacin- and rifampin-resistant strains and compared these with the same loci among 11 naturally occurring clinical recombinants to determine to what extent the properties of the *in vitro* model reflect natural recombination events.

MATERIALS AND METHODS

Reference strains and clinical isolates of *C. trachomatis*. The following 16 reference strains of *C. trachomatis* were used in this study: A/HAR-13, B/TW-5, D/UW-3, Da/TW-448, E/Bour, F/IC-Cal3, G/UW-57, H/UW-4, I/UW-12, Ia/UW-20, J/UW-36, Ja/UW-92, K/UW-31, L_1/440, L_2/434, and L_2b/UCH-1/proctitis. In addition, 11 urogenital samples were obtained from patients visiting sexually transmitted disease clinics. These samples represent *ompA* genotypes D, F, J, Ja, L_2, and L_2b and were identified as recombinants either by multilocus sequence typing (MLST) (9) or whole-genome sequencing (37). Briefly, both the reference and clinical strains were propagated in HeLa 229 cells, and the elementary bodies were harvested and purified from host cells via discontinuous density centrifugation in Renografin as we described previously (9, 37). The high-pure PCR template preparation kit (Roche Diagnostics) was used to extract genomic DNA following instructions provided by the manufacturer. The *ompA* genotype of the 16 reference strains and 11 clinical strains was confirmed by comparison with sequences in GenBank using BLAST (www.ncbi.nlm.nih.gov/BLAST/) as we described previously (36).

Selection and sequencing of specific genomic regions. Ten previously sequenced genomic regions of 16 clones of *in vitro*-derived recombinants of rifampin-resistant reference strain D (strain D:Rif^r-1, which has a mutation in the *rpoB* gene) and ofloxacin-resistant reference strain L_1 (strain L_1:Ofx^r-1, which has a mutation in the *gyrA* gene) (10) deposited in GenBank were analyzed ($n = 160$ sequences; 132,096 bp [8,256 per clone]). Similar regions including the *recF*, *incA*, *trpA*, *trpB*, *gyrA*, *rpoB*, *murA*, *rs2*, and *ompA* genes and intergenic regions (IGRs) between *rs2* and *ompA* and between *ompA* and *pbpB* were selected for sequencing the 11 clinical isolates ($n = 121$ sequences; 159,149 bp [14,454 to 14,481 per isolate]). PCR and sequencing primers (see Table S1 in the supplemental material) were designed based on available GenBank sequences of the respective genes for reference strain D/UW-3 (GenBank accession no. AE001273). Due to the large size of *rpoB* (3,759 bp) and *gyrA* (2,510 bp), sets of PCR primers were designed to generate overlapping amplicons for the entire gene. Four thermocycling profiles were used as shown in Table S2. The sequences of the clinical strains were subjected to BLAST and aligned to the corresponding gene or IGR of the 16 reference strains using MegAlign software (DNASTAR, Madison, WI) to identify single-nucleotide polymorphisms. The GenBank sequences for each *in vitro* clone were similarly analyzed in comparison to reference strains D/UW-3 and L_1/440. The least number of nucleotide mutations was used to determine the closest match to a strain for that gene or IGR.

Molecular evolution and phylogenetic analyses. For each of the *in vitro* and clinical recombinants, the percent G+C content of the genetic loci was calculated using EditSeq software (DNASTAR, Madison, WI). The molecular evolution of the genes of interest were each evaluated using the Nei-Gojobori method (29) to estimate the ratio of nonsynonymous to synonymous substitutions (d_N/d_S) as we have described previously (37). Using the *p*-distance model, values were normalized against the number of potential d_N and d_S sites; 95% confidence intervals were used to determine significant differences in the mean d_N and d_S. One thousand bootstrap replicates were performed to calculate the standard error.

The evolutionary history of the genes of interest was reconstructed using a bootstrap test of neighbor-joining tree topologies generated by MEGA (39) where trees were created using the Kimura 2-parameter model, and reconstructions at the protein level were created utilizing the gamma rate distribution model as we described previously (37).

Phylogenetic networks were constructed using SplitsTree4 (19) (http://www.splitstree.org/) as we described previously (9). The network estimates the phylogenetic relationship between all 16 reference strains and all 11 clinical strains and was drawn according to the default settings. The network was created based on 1,000 bootstrap replicates.

Determination of genomic crossover regions and breakpoints. Mosaic structures of the clinical isolates were proposed based on SimPlot analyses (http://sray.med.som.jhmi.edu/SCRoftware/simplot/) as we described previously (17). Briefly, the query sequence was compared with two potential parental strains and an outgroup strain using a window size of 100 bp with a 10-bp sliding step size to locate and determine breakpoints. Significant informative sites were determined using the maximum chi-squared test. Each informative site supports one phylogenetic tree based on the Kimura 2-parameter model with confidence levels for each branch in the tree calculated from 1,000 bootstrap replicates (17).

For each putative recombinant locus, a *P* value was calculated using Fisher's exact test for the best supported breakpoint region (i.e., between informative sites) based on the number of informative sites. A Bonferroni multiple test factor was conservatively applied for the number of ways to choose the observed number of breakpoints among the observed number of informative sites for the locus (this is a standard binomial coefficient). The resulting *P* value is the probability of observing the given pattern of informative sites supporting the two different ancestors in the absence of recombination.

Analysis of *in vitro* and clinical breakpoints. To compare the pattern of recombination among the clinical and *in vitro* sequences, a table of breakpoints was generated from the pooled clinical and *in vitro* data. First, data for seven genes (*recF*, *incA*, *trpA*, *gyrA*, *rpoB*, *murA*, and *ompA*) and the intervening regions (int) between each gene were arranged based on gene order in the genome. This analysis was performed using complete (clinical samples) and partial (*in vitro*) sequence data for the genes. The intervening regions do not represent sequence data; they represent only breakpoints between genes. Second, for each of the 27 samples, a value of 1 was entered if a recombination breakpoint was present within the gene and a value of 0 was entered if no breakpoint was observed. In each intervening region, if both genes before and after the region have the same strain, then a value of 0 was entered. Otherwise, a value of 1 was entered, indicating that a recombination breakpoint resides somewhere within the intervening region.

To determine whether the clinical and *in vitro* samples have different patterns of recombination, cluster analysis was performed. The patterns are different if the 27 samples can be grouped into distinct clinical and *in vitro* clusters based on the presence or absence of recombination breakpoints within each region. Before doing the cluster analysis, the binary data described above were converted to a dissimilarity matrix of Jaccard distances.

For each pair of samples A and B, the Jaccard dissimilarity is given by the equation $J' = (M_{01} + M_{10})/(M_{01} + M_{10} + M_{11})$ where M_{11} is the total number of regions where *A* and *B* both have a value of 1, M_{01} is the total number of regions where the value of *A* is 0 and the value of *B* is 1, and M_{10} is the total number of regions where the value of *A* is 1 and the value of *B* is 0.

The Jaccard dissimilarity matrix is calculated by the DISTANCE procedure in SAS. The output is used as the input to the CLUSTER procedure using Ward's minimum variance method (33a) where the distance between two clusters is the analysis of variance (ANOVA) sum of squares between the two clusters added up over all the samples. The output is then

displayed by using the TREE procedure. To determine the regions where the clusters have significantly different recombination patterns, a Wilcoxon score (rank sum) test from the NPAR1WAY procedure is then performed for each region from the output of the TREE procedure.

Nucleotide sequence accession numbers. The sequences from this study have been deposited in GenBank under the following accession numbers: JN795254 to JN795264 for *recF*, JN795281 to JN795291 for *incA*, JN795308 to JN795318 for *trpA*, JN795335 to JN795345 for *trpB*, JN795362 to JN795372 for *gyrA*, JN795389 to JN795399 for *rpoB*, JN795416 to JN795426 for *murA*, JN795470 to JN795480 for *rs2*, JN795443 to JN795453 for *ompA*, JN795497 to JN795507 for the IGR between *rs2* and *ompA*, and JN795524 to JN795534 for the IGR between *ompA* and *pbpB*.

RESULTS

Molecular evolution, phylogenetics, and recombination for the clinical recombinants. The initial analyses were based on the 11 genomic regions sequenced for 11 clinical recombinants in comparison with 16 reference strains. The d_N and d_S mutations and d_N/d_S were analyzed for each gene as shown in Fig. S1 in the supplemental material. All values except for IncA were <1, which would indicate an overall conservation of the protein structure with purifying selection for those loci. For IncA, however, a value of >1 suggests that it is under evolutionary pressure. There was a large variation for the mean genetic distance for the genes; the least variant gene was *trpB* (9.0 nucleotides [nt], ±0.1% nt, 6.0 amino acids [aa], ±0.3% aa) while the most variant was *ompA* (336 nt, ±11.2% nt, 94.0 aa, ±8.9% aa) (Fig. S2). *ompA* codes for the immunodominant major outer membrane protein of *C. trachomatis*. There was no significant variation in the G·C content for the genomic regions compared across *in vitro* and clinical recombinants (data not shown).

Phylogenetic reconstructions for the nine genes, excluding IGRs, were constructed based on the sequences for the 16 reference strains and 11 clinical strains (see Fig. S3 in the supplemental material). The trees were incongruent with each other, which supports prior evidence from the MLST data (9) that the clinical strains are recombinants.

The concatenated phylogenetic analysis of the genetic loci as shown in Fig. 1 revealed distinct clusters by invasive (lymphogranuloma venereum [LGV] strains, lavender), prevalent noninvasive (non-LGV strains, green), and less prevalent noninvasive (non-LGV strains, blue) urogenital disease grouping in which clinical Ja strains formed a distinct branch from the reference Ja strain (Fig. 1). The bootstrap SplitsTree decomposition result was congruent with the clustering in the MEGA phylogenetic tree and also displayed wide evolutionary paths, as indicated by the network structure between strains (Fig. 2), further supporting recombinant events in the evolution of the clinical strains.

Genomic crossover regions and breakpoints for *in vitro* and clinical recombinants. To determine putative crossover regions and breakpoints, the phylogenetic and BLAST analyses for each gene for the clinical samples were used to reveal putative mosaic structures. The right-hand column of Table S3 in the supplemental material displays the similarity for the closest hit to a reference strain for each clinical isolate. A graphical representation of the mosaics of the clinical recombinants is shown along with the *in vitro* recombinants in Fig. 3 based on sequence alignments (see Table S4 in the supplemental material for the genetic distance between clinical isolate and parental reference strains). Each colored arrow corresponds to a particular locus and the reference

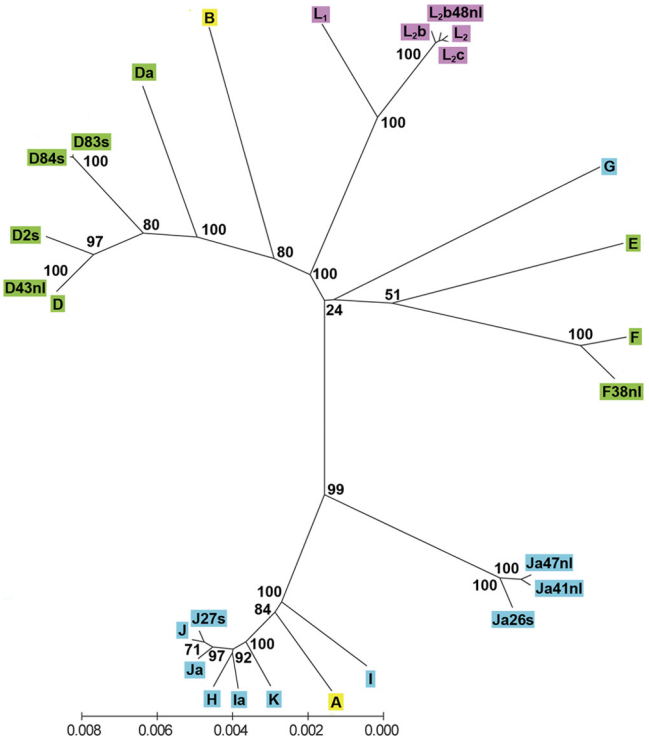

FIG 1 Phylogenetic tree for *C. trachomatis* strains. The 16 reference strains and 11 clinical strains are represented in the phylogeny. The tree was constructed with the concatenated sequences of the nine genes and two IGRs for the above strains using the neighbor-joining method in MEGA with 1,000 bootstrap replicates (see Materials and Methods). The numbers at tree branches indicate the percent bootstrap value as the level of confidence for the node. The colors refer to strains by disease groupings: lavender, invasive urogenital disease (LGV); blue, noninvasive urogenital disease; green, prevalent strains associated with noninvasive urogenital disease; and yellow, trachoma. The scale bar at the bottom of the figure represents the number of substitutions per site.

strain from which it is likely derived (see legend for Fig. 3). Ten of the 11 clinical isolates contain at least two breakpoint regions, as indicated by the black vertical arrows, which were confirmed by phylogenetic, SimPlot, and SAS analyses (see below). Open arrows indicate regions that could not be confirmed other than by comparative genetics. Interestingly, the clinical D43nl strain was identical to the reference strain D/UW-3 for the genomic regions sequenced, although it had previously been identified as a recombinant based on MLST where the sequences differed substantially in two housekeeping genes (9). Strains D2s, D83s, and D84s are recombinants of strains D, Da, and F with the exception of D2s where the parental strain is either H or Ia for the *rs2* gene. Strain F38nl is derived from strains E, F, and Da. Strain J27s is a recombinant of strains J and Ia, while the clinical Ja strains are mosaics of strains Da, E, F, and Ja.

Using SimPlot with phylogenetic, bootscan, and maximum chi-squared test analyses to identify the location of and confidence in breakpoint regions, each recombinant genomic region was confirmed. We used a 2×2 matrix for each putative breakpoint locus populated by the number of informative sites supporting the two ancestors on either side of the breakpoint. From this matrix, Fisher's exact test was calculated (Table 1). The total number of informative sites (t) and the number of breakpoint regions (c) used in

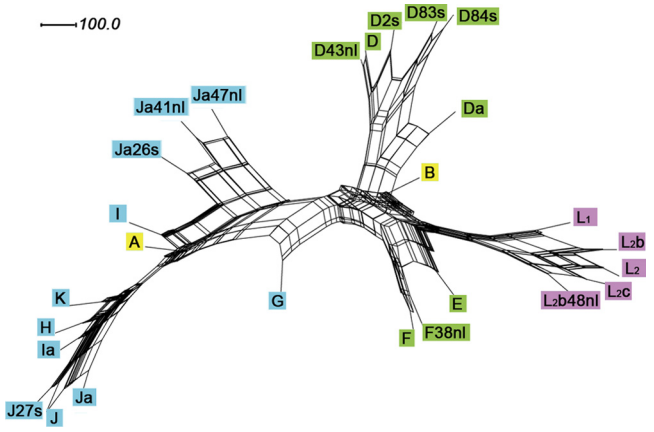

FIG 2 Phylogeny network interaction network for *C. trachomatis* strains. The tree was constructed using the concatenated sequences of the nine genes and two IGRs for the 16 reference strains and 11 clinical strains using SplitsTree (see Materials and Methods). Bootstrap replicates of 1,000 were performed to determine support for the tree. The colors refer to disease groupings: lavender, invasive urogenital disease; blue, noninvasive urogenital disease; green, prevalent strains causing urogenital disease; and yellow, trachoma. The scale bar represents the number of substitutions per site.

the analysis are shown. We applied Bonferroni's correction as the number of ways to distribute c in the intervals between t sites ($t - 1$ chooses c). Figure 4 illustrates the results for strains D83s, F38nl, and Ja41nl. The P values for the informative sites for the two breakpoints within the *rs2* region, the IGR between *rs2 and ompA*, and the *ompA* gene for strain D83s were determined as shown in Table 1. Breakpoints or breakpoint regions between arrows could not be determined because there was no contiguous sequence between the genes. The first crossover region likely occurs within *rs2* ($P = 6.2 \times 10^{-10}$), while the second is within the end of the IGR (flanked by *rs2* and *ompA*) and the beginning of *rs2* with a P value of 1.1×10^{-46}. For strain F38nl, the first breakpoint is within the end of the IGR (flanked by *rs2* and *ompA*) and the beginning of *rs2* ($P = 1.0$), while the second is within *ompA* with a P value of 1.1×10^{-60}. For strain Ja41nl, the crossover regions are the IGR flanked by *rs2* and *ompA* ($P = 1.0$) and the IGR flanked by *ompA* and *pbpB* ($P = 0.632$). The observed pattern for Ja41nl may have occurred by chance, based on our overly conservative P values, because there are not very many informative sites, but our findings are still consistent with the breakpoints we annotated.

These data further support the same two hot spots for clinical recombination identified in our previous work (17). Clinical strains D83s, D84s, and F38nl had a similar hot spot for recombination in the *rs2* and IGR (between *rs2* and *ompA*) genomic regions. In addition, clinical strains F38nl, Ja26s, Ja41nl, and Ja47nl had the same hot spot in the IGR (between *rs2* and *ompA*) and *ompA* regions.

Similar calculations were performed for the *in vitro*-derived sequences (data not shown). The nucleotide ranges within which the breakpoints occurred are shown in Table 2, and a graphical representation of the mosaics of the *in vitro* recombinants is shown in Fig. 3 based on sequence alignments.

We also calculated the P values for strain selection at each gene, using the same test each time (two-tailed Fisher's exact test), comparing the strain of the gene with the combined nearest neighbors on each side. We found that *trpA* (start [$P = 0.000001014$] and end [$P = 2.9 \times 10^{-9}$]), *gyrA* (start [$P = 4.3 \times 10^{-10}$]), *gyrB* (start [$P = 0.0000805$]), *rpoB* (start [$P = 4.3 \times 10^{-10}$]), and *ompA* ($P = 3.3 \times 10^{-8}$) were statistically significant but not *recF*, *ribF*, *incA*, *gyrA* (nucleotide position 237), *pmpC*, or *murA*. In analyzing a gene in this case, start refers to the beginning of the gene; the end (i.e., end of the gene) would be used as the neighbor unless the end and start agreed, in which case the next gene with data is used.

Comparison of *in vitro* and clinical breakpoints. The sequences of the 16 recombinant clones resulting from *in vitro* LGT of parental strains D:Rif^r-1 and L_1:Ofx^r-1 were analyzed for mutations using BLAST. Table 2 shows homology to the parental strain for each of the 10 genes sequenced, although only partial sequences were available in GenBank for each gene. The sequences of the 11 clinical samples were similarly analyzed for nine genes and two IGRs and are shown below the *in vitro* recombinants in Table 2. Precise breakpoints and the range (denoted by number in parentheses that are the base pair boundaries of the breakpoint regions) within which the breakpoint occurred could be identified in *trpA*, *gyrA*, and *rpoB* for most of the *in vitro* clones (Table 2) and in *incA*, *trpB* or *trpA*, *rs2*, the *rs2* IGR, *ompA*, and the *ompA* IGR for the clinical recombinants (Table 1).

The Ward's minimum variance clustering procedure separated the samples into clinical and *in vitro* clusters without any misclassification (Fig. 5). For the two clusters, the squared multiple correlation R^2 (the proportion of variance accounted for by the two clusters) was 0.32. As shown in the tree (Fig. 5), the clinical cluster can be further subdivided into two branches. These two clusters both contain samples of clinical D and L strains.

The Wilcoxon rank sum test (see Table S5 in the supplemental material) resulted in a P value of < 0.0001 for *trpA*, intragene region 4 (int4), *gyrA*, and int5, while the P value was 0.02 for *rpoB*. All other comparisons were insignificant, showing that this genomic region is where the recombination pattern is different between clinical and *in vitro* recombinants.

DISCUSSION

There is a growing body of research that supports naturally occurring intra- and interspecies recombination for numerous *Chlamydia* spp. (13, 14, 17, 18, 25, 30, 31). Despite an abundance of gene sequences and a number of whole-genome sequences, the mechanism(s) of LGT remains elusive. Without this knowledge, the development of a stable gene transfer system has been greatly hampered. In addition, *C. trachomatis* has been largely refractory to directed mutagenesis, although a recent study by Wang et al. (40) showed that plasmid-free *C. trachomatis* can be transfected with a chlamydial plasmid to restore glycogen synthesis. Thus, the current lack of reliable methods for genetic manipulation of *C. trachomatis* profoundly limits our ability to understand the function of many genes involved in virulence and host specificity.

In vitro studies of antibiotic-resistant strains have indeed yielded recombinant organisms that may be useful for evaluating recombination mechanisms (10, 11, 38). However, the results remain difficult to interpret, because while some clones have identical breakpoint regions (e.g., in *gyrA*), there are no common sites of crossover among recombinants produced by the same or different research groups using the same resistance markers (10, 11, 38). For example, the L_1:Ofx^r-1 and L_1:Rif^r-1 strains that were crossed in one study (11) resulted in recombinant regions in *gyrA* and *rpoB* that differed from recombinant regions of the D:Rif^r-1 and L_1:Ofx^r-1 crosses in a second study by the same group (10).

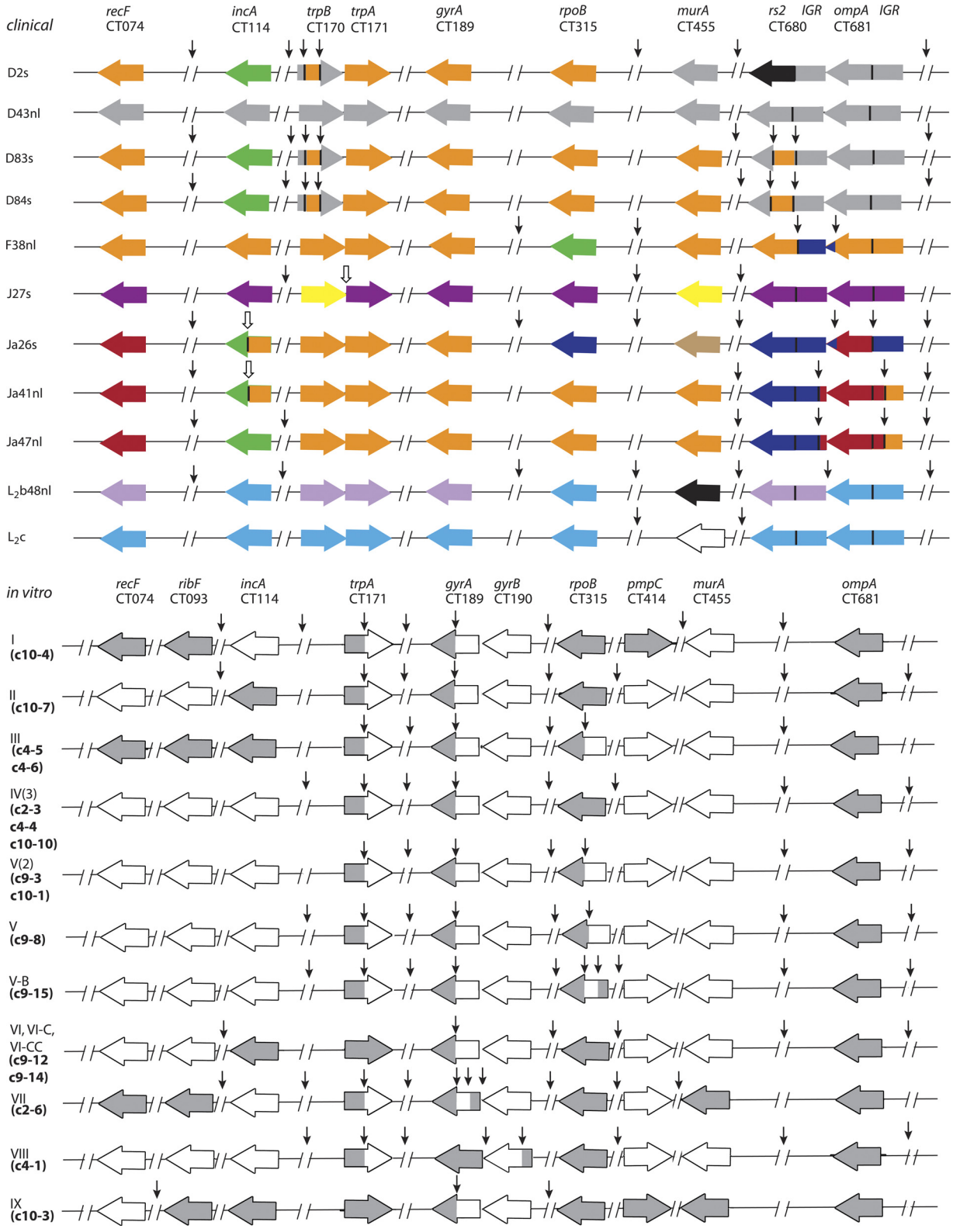

FIG 3 Proposed chromosomal mosaic structure for the 11 clinical recombinants (top half) and the 16 *in vitro* recombinants (bottom half) of *C. trachomatis*. The genomic loci are represented by horizontal colored arrows, where the arrow direction signifies the coding strand direction of the gene. The color of the arrow denotes the reference strain from which the mosaic is most likely derived as follows: gray, strain D; green, strain Da; dark blue, strain E; orange, strain F; beige, strain I; yellow, strain Ia; purple, strain J; red, strain Ja; white, strain L_1; light blue, strain L_2; and lavender, strain L_2b. The black horizontal arrows for strains D2s and L_2b48nl denote ambiguity, as the parental strains may be H/UW-4 or Ia/UW-202 for *rs2* or D/UW-3, Ia/UW-20, or K/UW-31 for *murA*, respectively. The boxes for the clinical recombinants represent the two noncoding IGRs, between *rs2* and *ompA*, and upstream of *ompA*. The proposed crossover regions are indicated by vertical arrows determined by sequence alignments and statistical analyses (see Materials and Methods). The open arrows were confirmed by sequence alignments only. The roman numeral notation for the *in vitro* recombinant strains corresponds to the identification tags of the DeMars and Weinfurter (10) recombinants. Labels indicated in bold type and parentheses [e.g., (c10–4)] are the corresponding names of the *in vitro* clones in GenBank.

TABLE 1 *P* value calculations for clinical recombinants based on SimPlot analyses

Clinical isolate[a]	n11[b]	n12	n21	n22	No. of informative sites (outgroup)	Crossovers (location)[c]	Fisher's exact test *P* value	No. of tests[d]	Bonferroni-corrected *P* value
D2s (L)	2	0	0	1	64 (B)	2 (*trpB*, positions 270 to 320)	0.33	2,016	1.0
D2s (R)	0	1	60	0	64 (B)	2 (*trpB*, position 565)	0.02	2,016	1.0
D83s (L)	2	0	0	37	91 (A)	2 (*trpB*, position 324)	0.00135	4,005	0.995
D83s (R)	0	37	52	0	91 (A)	2 (*trpB*, positions 848 to 886)	6.73×10^{-26}	4,005	2.69×10^{-22}
D83s (L)	20	0	0	33	500 (L_2)	2 (*rs2*, position 525)	4.94×10^{-15}	124,251	6.2×10^{-10}
D83s (R)	0	33	447	0	500 (L_2)	2 (IGR [between *rs2* and *ompA*], positions 365 to 372, to *rs2*, positions 1 to 11)	8.84×10^{-52}	124,251	1.1×10^{-46}
D84s (L)	2	0	0	37	91 (A)	2 (*trpB*, position 324)	0.00135	4,005	0.995
D84s (R)	0	37	52	0	91 (A)	2 (*trpB*, positions 848 to 886)	6.73×10^{-26}	4,005	2.69×10^{-22}
D84s (L)	20	0	0	33	122 (L_2)	2 (*rs2*, position 525)	4.94×10^{-15}	7,260	3.6×10^{-11}
D84s (R)	0	33	69	0	122 (L_2)	2 (IGR [between *rs2* and *ompA*], positions 343 to 372, to *rs2*, positions 1 to 2)	1.55×10^{-27}	7,260	1.1×10^{-23}
F38nl (L)	2	0	0	55	350 (I)	2 (IGR [between *rs2* and *ompA*], positions 346 to 377, to *rs2*, positions 1 to 89)	0.00062657	60,726	1.0
F38nl (R)	0	55	293	0	350 (I)	2 (*ompA*, positions 926 to 1082)	1.88×10^{-65}	60,726	1.1×10^{-60}
J27s					10 (A)	1 (*trpB* or *trpA*, no breakpoint detected)			
Ja26s					None (L_2)	1 (*incA*, no breakpoint detected)			
Ja26s (L)	38	0	0	149	190 (L_2)	2 (*ompA*, positions 1016 to 1186)	1.39×10^{-40}	17,766	2.584^{-38}
Ja26s (R)	0	149	3	0	190 (L_2)	2 (IGR [between *ompA* and *pbpB*], positions 447 to 608, to *ompA*, positions 1 to 51)	0.00000174	17,766	0.030
Ja41nl					None (L_2)	1 (*incA*, no breakpoint detected)			
Ja41nl (L)	1	0	0	105	108[e]	2 (*ompA*, positions 1071 to 1117)	0.00943	5,671	1.0
Ja41nl (R)	0	105	2	0	108[e]	2 (IGR [between *ompA* and *pbpB*], positions 518 to 551)	1.76×10^{-4}	5,671	0.632
Ja47nl (L)	1	0	0	104	107[e]	2 (*ompA*, positions 1093 to 1191, to IGR [between *rs2* and *ompA*], positions 1 to 377)	0.00952	5,565	1.0
Ja47nl (R)	0	104	3	0	107[e]	2 (IGR [between *ompA* and *pbpB*], positions 548 to 586, to *ompA*, positions 1 to 123)	5.04×10^{-06}	5,565	0.038
L_2b48nl	1	0	0	1	2 (K)	1 (IGR [between *rs2* and *ompA*], positions 48 to 151)	1	1	1

[a] The letter in parentheses after the clinical isolate designation indicates the location of the crossover: L for crossover on the left and R for crossover on the right (for strains D83s, F38nl, and Ja41nl; see Fig. 4).
[b] n11 through n22 refer to entries in a 2 × 2 contingency table; for example, there are xx informative sites in the first part of the region that favor one tree, and none that favor the other tree, so the first column of the 2 × 2 contingency table has a xx and a 0.
[c] The numbers in parentheses refer to the location in the gene based on the start codon of the reference strain for that clinical strain.
[d] Number of tests used to determine the Bonferroni-corrected *P* value. Bonferroni's correction (Sidak formula) was applied because the regions of sequence used to define the contingency tables were chosen based on the data, not in advance. Because of correlations in the multiple test hypotheses, this correction is overly conservative.
[e] Because the parental strains were Ja, E, and F, no outgroup could be used given the limitations of SimPlot.

In the present study, we examined multiple genomic loci for 16 *in vitro*-derived recombinants from strains D:Rif^r-1 and L_1:Ofx^r-1 and compared these with the same loci from 11 clinical recombinants. We initially determined the phylogenies and then the precise regions within which breakpoints occur for nine genes and two IGRs among the 11 clinical strains that were previously identified as recombinants based on MLST (9). Up to five statistically significant breakpoint regions were identified in areas of contiguous gene or IGR sequences among nine clinical recombinants (Fig. 3 and Tables 1 and 2). Three regions, including two hot spots for recombination involving *rs2*, the IGR (between *rs2* and *ompA*), and *ompA*, had previously been identified among *C. trachomatis* strains from a patient population from Lisbon, Portugal, that we had studied (17). The samples in the present study were from San Francisco, California, and Amsterdam, The Netherlands, suggesting that not only are these regions common sites of recombination in the genome, but they are unaffected by geographic perturbations. We and others have also shown the *ompA* region to be involved in homologous recombination (5, 16, 17, 20–22, 25) in addition to the *trpB* or *trpA* region (20). Only one strain in the current study, D43nl, showed no evidence of recombination despite previous identification of LGT in our MLST study (9). This is not a contradiction, given that the MLST scheme produces sequences of seven housekeeping genes not represented among the genomic regions sequenced in this study. In addition, the D43nl strain was from The Netherlands, while the other clinical D strains originated in the San Francisco Bay area in California, suggesting that there are some geographic differences within the strains. This is not caused by a lack of genome diversity among strains in The Netherlands because we previously found strain diversity in that country (9). The exact recombinant nature of strains such as D43nl would be revealed by whole-genome sequencing, which

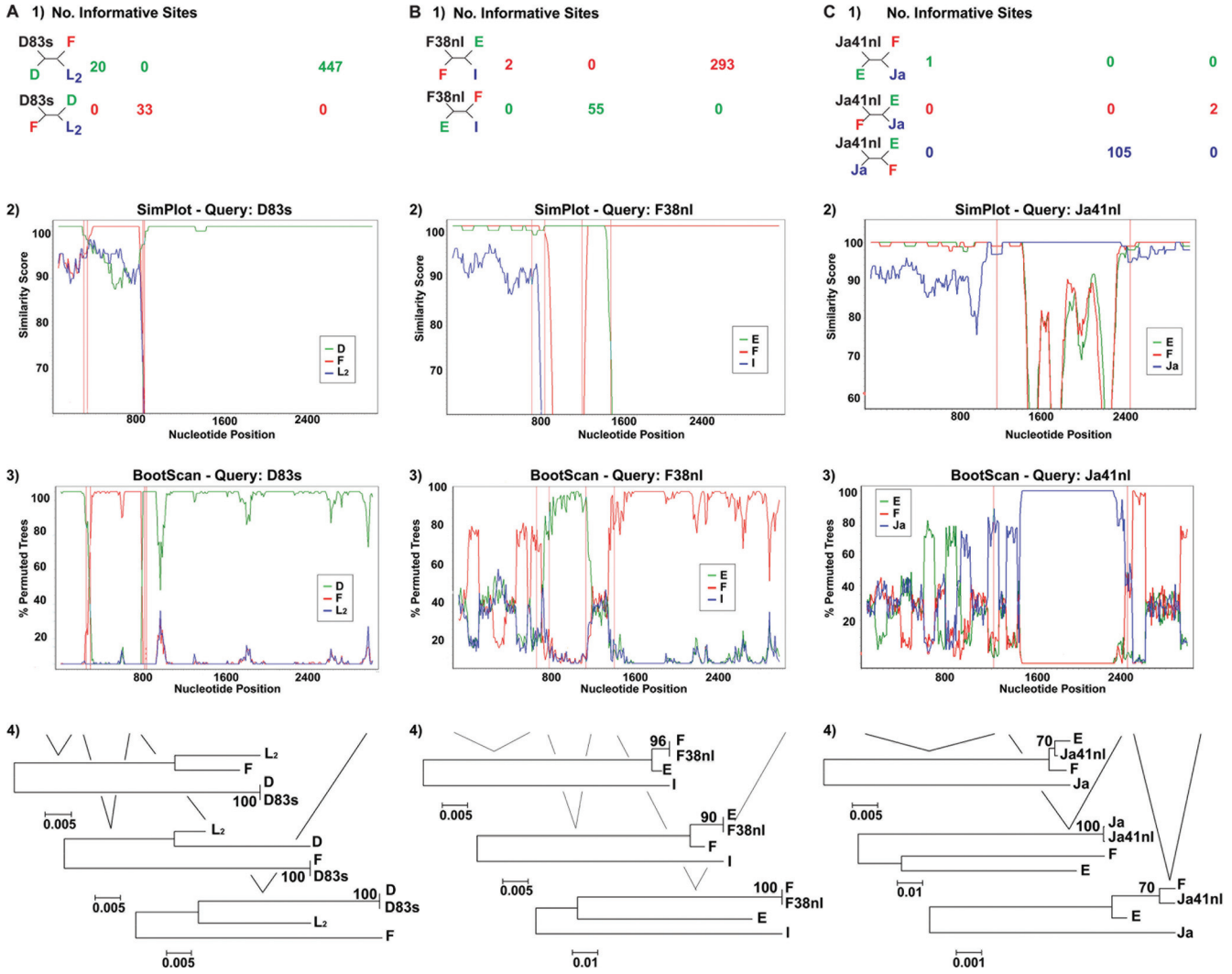

FIG 4 SimPlot representation of the proposed crossover regions for clinical isolates D83s (A), F38nl (B), and Ja41nl (C). The vertical lines in panels 2 and 3 correspond to the approximate breakpoint, while the nucleotide positions at the bottom of each figure correspond to alignment positions of the region analyzed. The genomic regions shown are from *rs2* to *ompA* for isolate D83s, from *rs2* to *ompA* for isolate F38nl, and from *rs2* to the IGR upstream of *ompA* for isolate Ja41nl (panels 1). The number of informative sites shared by the recombinant sequence (black) and the parental sequence from which the mosaic is likely derived (red, green, or blue) are shown. Informative sites and supporting four-member trees were determined for each region bounded by a breakpoint (panels 2). SimPlot analysis depicting the similarity of the recombinant sequences (query) to their respective parental reference strains in a 100-bp window with a step size of 10 bp (panels 3). BootScanning analysis showing the phylogenetic relatedness between these sequences (panels 4). Neighbor-joining (NJ) trees constructed for each region bounded by a breakpoint. NJ trees based on the Kimura 2-parameter model and BootScan confidence levels were calculated for 1,000 replicates.

should become standard protocol once costs reach a reasonable level for research laboratories.

We also evaluated the 16 *in vitro* recombinants for statistically significant breakpoint regions (Table 2). The majority of breakpoints occurred within *gyrA* and *rpoB*; 14 contained identical breakpoint regions in *gyrA* with one recombinant showing an additional breakpoint region in *gyrA*, while *rpoB* contained variable breakpoint regions for six recombinants. The majority of the *gyrA* sequence was L_1 for the 14 crosses and D for the six *rpoB* crosses.

Surprisingly, the *in vitro* recombinant VIII had an entirely D sequence for *gyrA*, which should have been selected against, since strains D:Rif^r-1 and L_1:Ofx^r-1 were used in the cross. Ofloxacin is one of many fluoroquinolone antibiotics, and previous studies of fluoroquinolone-resistant isolates revealed single-nucleotide polymorphisms conferring resistance within the established quinolone resistance-determining region of *gyrA* but no mutations in the *gyrB*, *parC*, or *parE* quinolone resistance-determining regions (12, 32). However, one study found that isolates with elevated MICs against fluoroquinolones did not have *gyrA* or *parC* mutations in positions known to be critical for fluoroquinolone resistance (43), suggesting that other genes may be involved in resistance. In our study, we identified a mutation in *gyrB* that was the breakpoint crossover from D to L_1 where the majority of the *gyrB* sequence was L_1. Consequently, *gyrB* may be able to confer fluoroquinolone resistance, as this is the most plausible explanation for the occurrence of the VIII clone.

In comparing all 27 recombinants, two distinct clusters were identified, and the two clusters reflected the clinical and *in vitro* groups (Fig. 5), indicating a significant difference in occurrence of recombination breakpoints within the genes we sequenced. Be-

TABLE 2 Genomic crossover regions and specific breakpoints for *in vitro* and clinical recombinants[f]

		Genes													
In vitro[a]	Clone[b]	*recF* CT074	*ribF* CT093	*incA* CT119	*trpB* CT170	*trpA*[c] CT171	*gyrA*[c] CT189	*gyrB*[c] CT190	*rpoB*[c] CT315	*pmpC* CT414	*murA* CT455	*rs2* CT680	IGR	*ompA* CT681	IGR
I	C10-4	D	D	L_1		(D)344-477(L_1)	(D)104-237(L_1)	L_1	D	D	L_1			D	
II	C10-7	L_1	L_1	D		(D)344-477(L_1)	(D)104-237(L_1)	L_1	D	L_1	L_1			D	
III (2)	C4-5, C4-6	D	D	D		(D)344-477(L_1)	(D)104-237(L_1)	L_1	(D)1044-1560(L_1)[d]	L_1	L_1			D	
IV (3)	C2-3, C4-4, C10-10	L_1	L_1	L_1		(D)344-477(L_1)	(D)104-237(L_1)	L_1	D	L_1	L_1			D	
V (2)	C9-3, C10-1	L_1	L_1	L_1		(D)344-477(L_1)	(D)104-237(L_1)	L_1	(D)1044-1560(L_1)[d]	L_1	L_1			D	
V-A	C9-8	L_1	L_1	L_1		(D)344-477(L_1)	(D)104-237(L_1)	L_1	(D)1560-1714(L_1)	L_1	L_1			D	
V-B	C9-15	L_1	L_1	L_1		(D)344-477(L_1)	(D)104-237(L_1)	L_1	(D)1714-1953(L_1), (L_1)1953-2226(D)	L_1	L_1			D	
VI	C9-12	L_1	L_1	D		D	(D)104-237(L_1)	L_1	D	L_1	L_1			D	
VI-C, VI-CC	C9-14	L_1	L_1	D		D	(D)104-237(L_1)	L_1	D	L_1	L_1			D	
VII	C2-6	D	D	L_1		(D)344-477(L_1)	(D)104-237(L_1), (L_1)237-492(D)[e]	L_1	D	L_1	D			D	
VIII	C4-1	L_1	L_1	L_1		(D)344-477(L_1)	D	(L_1)2004-2371(D)	D	L_1	L_1			D	
IX	C10-3	L_1	D	D		D	(D)104-237(L_1)	L_1	D	D	D			D	

Clinical	ID no	*recF* CT074	*ribF* CT093	*incA* CT119	*trpB* CT170	*trpA*[c] CT171	*gyrA*[c] CT189	*gyrB*[c] CT190	*rpoB*[c] CT315	*pmpC* CT414	*murA* CT455	*rs2* CT680	IGR	*ompA* CT681	IGR
	D2s	F		Da	D F D	F	F		F		D	H/Ia	D	D	D
	D43nl	D		D	D	D	D		D		D	D	D	D	D
	D83s	F		Da	D F D	F	F		F		F	D F	F D	D	D
	D84s	F		Da	D F D	F	F		F		F	D F	F D	D	D
	F38nl	F		F	F	F	F		Da		F	F	E	E F	F
	J27s	J		J	Ia	J	J		J		Ia	J	J	J	J
	Ja26s	Ja		Da F	F	F	F		E		I	E	E	E Ja E	E
	Ja41nl	Ja		Da F	F	F	F		F		F	E	E Ja	Ja	Ja F
	Ja47nl	Ja		Da	F	F	F		F		F	E	E Ja	Ja	Ja F
	L_2b48nl	L_2b		L_2	L_2b	L_2b	L_2b		L_2		Da/I/K	L_2b	L_2b	L_2	L_2
	L_2c	L_2		L_2	L_2	L_2	L_2		L_2		L_1	L_2	L_2	L_2	L_2

[a] Sixteen recombinant clones resulting from *in vitro* LGT (10).

[b] Clone sequences were retrieved from the GenBank database under the following accession numbers for the indicated genes: EU104987 to EU105002 for *gyrA*, EU105003 to EU105018 for *gyrB*, EU105019 to EU105034 for *incA*, EU105035 to EU105050 for *murA*, EU105051 to EU105066 for *ompA*, EU105067 to EU105082 for *pmpC*, EU105083 to EU105098 for *recF*, EU105099 to EU105114 for *ribF*, EU105115 to EU105130 for *rpoB*, and EU105131 to EU105146 for *trpA*. For the clinical strains, the identification number (ID no) is shown.

[c] Nucleotide positions of crossover regions.

[d] Spontaneous mutation occurring at nucleotide 1400.

[e] Spontaneous mutation occurring at nucleotide 249.

[f] The source of the genes is indicated by color as follows: gray, strain D; olive green, strain Da; dark blue, strain E; pumpkin-brown, strain F; beige, strain I; yellow, strain Ia; purple, strain J; red-brown, strain Ja; white, strain L_1; light blue, strain L_2; and lavender, strain L_2b.

cause the clinical cluster contained D and L strains in two different branches, neither D nor L have patterns of recombination that are distinct from other strains. Therefore, the clustering is not due to the subtypes used in the *in vitro* experiments (D and L) being different from other strains. The *in vitro* recombinants had significantly more breakpoints in *gyrA*, which confers ofloxacin resistance, and in *rpoB*, which confers rifampin resistance, than the clinical recombinants. This was not surprising due to the introduced antibiotic pressure that selects for recombination in genes that confer a specific resistance, although the crossover breakpoints could have occurred outside the gene boundaries.

The *trpA* gene also showed significantly more breakpoints with

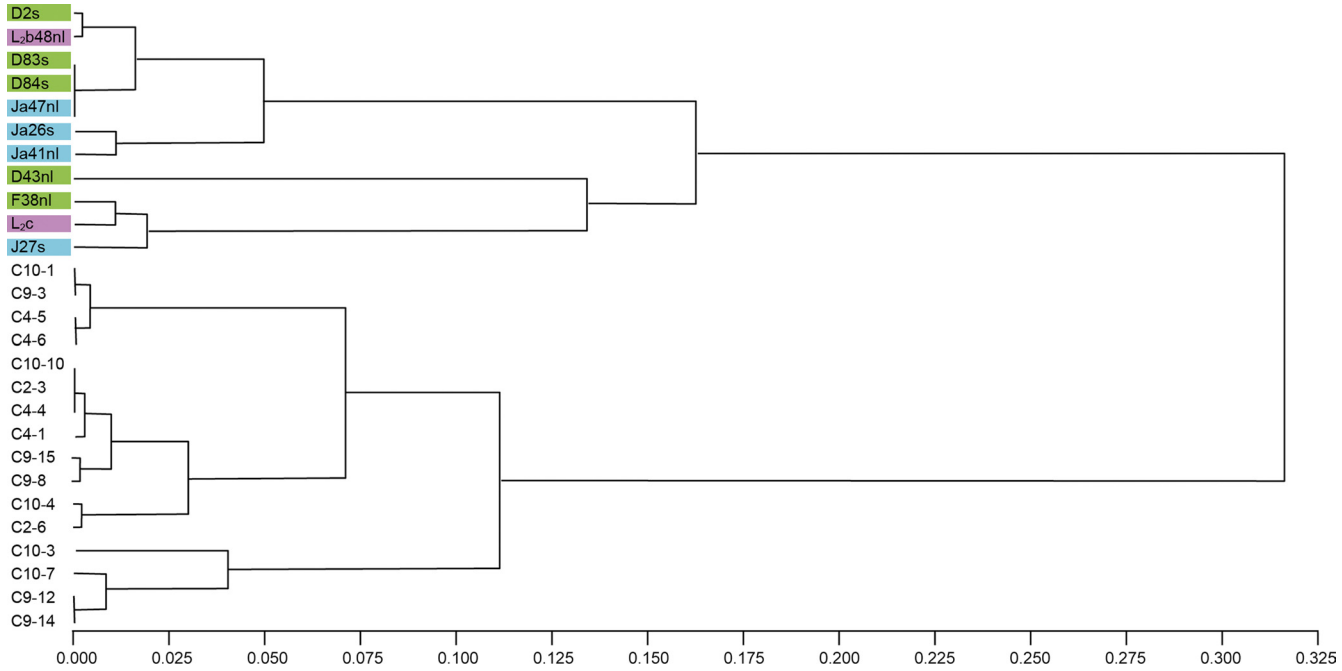

FIG 5 Phylogenetic tree showing the hierarchical relationship between the *in vitro* and clinical recombinants calculated using the Jaccard coefficient as a distance metric (see Materials and Methods). The length of each branch represents the similarity in the pattern of breakpoint locations.

an identical breakpoint region for 13 (81.25%) of the 16 *in vitro* recombinants that contained a crossover in this gene (Table 2). This was unexpected because rifampin and ofloxacin have not been known to induce mutations in *trpA*, although there may be some selection pressure within, upstream, or downstream of this gene. *trpA* encodes TrpA, an alpha chain that interacts with another alpha chain and two TrpB beta chains to form the tetrameric tryptophan synthase. Tryptophan is an essential amino acid for replication during differentiation of the metabolically active reticulate body into the infectious elementary body (4) that, upon release from the cell, sets up another round of infection in adjacent tissue or infects a new host following mucosal transmission. TrpR, TrpB, and TrpA are products of genes that make up the partial tryptophan operon, which is functional in the sexually transmitted strains of *C. trachomatis* but truncated and rendered nonfunctional among ocular trachoma strains (1, 6). In *in vitro* cell culture systems, gamma interferon (IFN-γ), which has been observed to be endogenously produced by *C. trachomatis* infection of HeLa cells (7), induces expression of indoleamine 2,3-dioxygenase that in turn degrades tryptophan. The inhibitory effect of IFN-γ varies according to strain. It is less effective against the LGV strains, which include L_1, compared with other urogenital strains; this effect is likely linked to two amino acid polymorphisms at residues 177 and 178 in the C terminus of TrpA (6). Each of the 13 *in vitro* TrpA recombinants were derived from L_1 in the C-terminal half of the protein, a critical region of loop 6 involved in alpha-beta subunit interactions and binding of important substrates such as indole-3-glycerol phosphate (24, 41). Consequently, while IFN-γ was not added to the growth medium, it is probable that *in vitro* induced recombination that places L_1 in the C terminus of TrpA contributes to activity of the protein and counters the effects of endogenously produced IFN-γ (7) in a way that is advantageous to the progeny.

There was also significant selection among the *in vitro* recombinants for *ompA* ($P = 3.3 \times 10^{-8}$), which, like *trpA*, is not known to be involved in antibiotic resistance. The strongest bias was for strain D (Table 2). Our previous work on clinical recombinants revealed two breakpoint regions, one of which was identified as a hot spot that flanks the *ompA* gene (17). Our findings support these data. The similarities in breakpoint regions flanking *ompA* among the *in vitro* and clinical recombinants in this study suggest a common mechanism of LGT, but additional studies will be needed to identify the precise mechanisms involved.

Recombination is prevalent in *C. trachomatis* and plays a major role in its genetic diversification and evolution. Our findings suggest that *in vitro*-derived recombinants of antibiotic resistance strains do not reflect the natural statistical distribution of recombination events. While frequent LGT with multiple breakpoints may occur *in vivo* among various strains, these events may be selected against given the mucosal environment or may produce mutants that are less fit. In any case, the current *in vitro* model should be used to evaluate mechanism(s) of recombination given the observed selection biases at the unexpected loci of *trpA* and *ompA*. Additional studies are needed to determine the rate and diversity of recombination for *C. trachomatis* clinical samples. In this process, comparative and functional genomics of many clinical and *in vitro* isolates that do not rely on antibiotic selection will be key to determining the precise mechanism of LGT in *C. trachomatis* and, ultimately, to develop a robust gene transfer system for the organism.

ACKNOWLEDGMENTS

This research was supported in part by Public Health Service grant R01 AI059647 (to D.D.) and by National Science Foundation-U.S. Department of Agriculture grant 2009-65109-05760 (to D.D.).

We thank Michael Landis for excellent technical assistance and Ser-

vaas Morre and Henry de Vries for allowing us to use clinical samples from The Netherlands.

REFERENCES

1. **Akers JC, Tan M.** 2006. Molecular mechanism of tryptophan-dependent transcriptional regulation in *Chlamydia trachomatis.* J. Bacteriol. **188**: 4236–4243.
2. **Bhengraj AR, Vardhan H, Srivastava P, Salhan S, Mittal A.** 2010. Decreased susceptibility to azithromycin and doxycycline in clinical isolates of *Chlamydia trachomatis* obtained from recurrently infected female patients in India. Chemotherapy **56**:371–377.
3. **Binet R, Maurelli AT.** 2007. Frequency of development and associated physiological cost of azithromycin resistance in *Chlamydia psittaci* 6BC and *C. trachomatis* L2. Antimicrob. Agents Chemother. **51**:4267–4275.
4. **Boehm U, Klamp T, Groot M, Howard JC.** 1997. Cellular responses to interferon-gamma. Annu. Rev. Immunol. **15**:749–795.
5. **Brunelle BW, Sensabaugh GF.** 2006. The *ompA* gene in *Chlamydia trachomatis* differs in phylogeny and rate of evolution from other regions of the genome. Infect. Immun. **74**:578–585.
6. **Caldwell HD, et al.** 2003. Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. J. Clin. Invest. **111**:1757–1769.
7. **Cantell K.** 1961. Production and action of interferon in HeLa cells. Arch. Gesamte Virusforsch. **10**:510–521.
8. **Centers for Disease Control and Prevention.** 2010. Sexually transmitted diseases treatment guidelines: chlamydial infections.Centers for Disease Control and Prevention, Atlanta, GA.
9. **Dean D, et al.** 2009. Predicting phenotype and emerging strains among *Chlamydia trachomatis* infections. Emerg. Infect. Dis. **15**:1385–1394.
10. **DeMars R, Weinfurter J.** 2008. Interstrain gene transfer in *Chlamydia trachomatis* in vitro: mechanism and significance. J. Bacteriol. **190**: 1605–1614.
11. **DeMars R, Weinfurter J, Guex E, Lin J, Potucek Y.** 2007. Lateral gene transfer in vitro in the intracellular pathogen *Chlamydia trachomatis.* J. Bacteriol. **189**:991–1003.
12. **Dessus-Babus S, Bebear CM, Charron A, Bebear C, de Barbeyrac B.** 1998. Sequencing of gyrase and topoisomerase IV quinolone-resistance-determining regions of *Chlamydia trachomatis* and characterization of quinolone-resistant mutants obtained in vitro. Antimicrob. Agents Chemother. **42**:2474–2481.
13. **Dugan J, Rockey DD, Jones L, Andersen AA.** 2004. Tetracycline resistance in *Chlamydia suis* mediated by genomic islands inserted into the chlamydial *inv*-like gene. Antimicrob. Agents Chemother. **48**:3989–3995.
14. **Fitch WM, Peterson EM, de la Maza LM.** 1993. Phylogenetic analysis of the outer-membrane-protein genes of *Chlamydia*, and its implication for vaccine development. Mol. Biol. Evol. **10**:892–913.
15. **Gerrish PJ, Colato A, Perelson AS, Sniegowski PD.** 2007. Complete genetic linkage can subvert natural selection. Proc. Natl. Acad. Sci. U. S. A. **104**:6266–6271.
16. **Gomes JP, Bruno WJ, Borrego MJ, Dean D.** 2004. Recombination in the genome of *Chlamydia trachomatis* involving the polymorphic membrane protein C gene relative to *ompA* and evidence for horizontal gene transfer. J. Bacteriol. **186**:4295–4306.
17. **Gomes JP, et al.** 2007. Evolution of *Chlamydia trachomatis* diversity occurs by widespread interstrain recombination involving hotspots. Genome Res. **17**:50–60.
18. **Gomes JP, et al.** 2006. Polymorphisms in the nine polymorphic membrane proteins of *Chlamydia trachomatis* across all serovars: evidence for serovar Da recombination and correlation with tissue tropism. J. Bacteriol. **188**:275–286.
19. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23**:254–267.
20. **Jeffrey BM, et al.** 2010. Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. Infect. Immun. **78**:2544–2553.
21. **Joseph SJ, Didelot X, Gandhi K, Dean D, Read TD.** 2011. Interplay of recombination and selection in the genomes of *Chlamydia trachomatis.* Biol. Direct **6**:28.
22. **Lysen M, et al.** 2004. Characterization of *ompA* genotypes by sequence analysis of DNA from all detected cases of *Chlamydia trachomatis* infections during 1 year of contact tracing in a Swedish county. J. Clin. Microbiol. **42**:1641–1647.
23. **Manavi K.** 2006. A review on infection with *Chlamydia trachomatis.* Best Pract. Res. Clin. Obstet. Gynaecol. **20**:941–951.
24. **Miles EW.** 2001. Tryptophan synthase: a multienzyme complex with an intramolecular tunnel. Chem. Rec. **1**:140–151.
25. **Millman K, Tavaré S, Dean D.** 2001. Recombination in the *ompA* gene but not the *omcB* gene of *Chlamydia* contributes to serovar-specific differences in tissue tropism, immune surveillance, and persistence of the organism. J. Bacteriol. **183**:5997–6008.
26. **Misyurina OY, et al.** 2004. Mutations in a 23S rRNA gene of *Chlamydia trachomatis* associated with resistance to macrolides. Antimicrob. Agents Chemother. **48**:1347–1349.
27. **Muller HJ.** 1964. The relation of recombination to mutational advance. Mutat. Res. **106**:2–9.
28. **Muller HJ.** 1932. Some genetic aspects of sex. Am. Nat. **66**:118–138.
29. **Nei M, Kumar S.** 2000. Molecular evolution and phylogenetics. Oxford University Press, New York, NY.
30. **Read TD, et al.** 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Res. **28**:1397–1406.
31. **Read TD, et al.** 2003. Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae. Nucleic Acids Res. **31**:2134–2147.
32. **Rupp J, Solbach W, Gieffers J.** 2008. Variation in the mutation frequency determining quinolone resistance in *Chlamydia trachomatis* serovars L2 and D. J. Antimicrob. Chemother. **61**:91–94.
33. **Samra Z, Rosenberg S, Soffer Y, Dan M.** 2001. *In vitro* susceptibility of recent clinical isolates of *Chlamydia trachomatis* to macrolides and tetracyclines. Diagn. Microbiol. Infect. Dis. **39**:177–179.

33a. **SAS Institute Inc.** 2004. SAS/STAT 9.1 user's guide. SAS Institute Inc, Cary, NC.

34. **Shkarupeta MM, Lazarev VN, Akopian TA, Afrikanova TS, Govorun VM.** 2007. Analysis of antibiotic resistance markers in *Chlamydia trachomatis* clinical isolates obtained after ineffective antibiotic therapy. Bull. Exp. Biol. Med. **143**:713–717.
35. **Somani J, Bhullar VB, Workowski KA, Farshy CE, Black CM.** 2000. Multiple drug-resistant *Chlamydia trachomatis* associated with clinical treatment failure. J. Infect. Dis. **181**:1421–1427.
36. **Somboonna N, Mead S, Liu J, Dean D.** 2008. Discovering and differentiating new and emerging clonal populations of *Chlamydia trachomatis* with a novel shotgun cell culture harvest assay. Emerg. Infect. Dis. **14**: 445–453.
37. **Somboonna N, et al.** 2011. Hypervirulent *Chlamydia trachomatis* clinical strain is a recombinant between lymphogranuloma venereum (L2) and D lineages. mBio **2**(3):e00045–11. doi:10.1128/mBio.00045-11.
38. **Suchland RJ, Sandoz KM, Jeffrey BM, Stamm WE, Rockey DD.** 2009. Horizontal transfer of tetracycline resistance among *Chlamydia* spp. in vitro. Antimicrob. Agents Chemother. **53**:4604–4611.
39. **Tamura K, et al.** 2011. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
40. **Wang Y, et al.** 2011. Development of a transformation system for *Chlamydia trachomatis*: restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector. PLoS Pathog. **7**:e1002258.
41. **Weyand M, Schlichting I.** 1999. Crystal structure of wild-type tryptophan synthase complexed with the natural substrate indole-3-glycerol phosphate. Biochemistry (Mosc.) **38**:16469–16480.
42. **World Health Organization.** 2001. Global prevalence and incidence of selected curable sexually transmitted infections: overview and estimates. WHO reference numbers WHO/HIV_AIDS/2001.02 and WHO/CDS/CSR/EDC/2001/10.Department of HIV/AIDS, World Health Organization, Geneva, Switzerland.
43. **Yokoi S, et al.** 2004. Uncommon occurrence of fluoroquinolone resistance-associated alterations in GyrA and ParC in clinical strains of *Chlamydia trachomatis.* J. Infect. Chemother. **10**:262–267.